三维重建系统下的特征点处理与位姿恢复优化算法 *

徐建鹏，卜凡亮†

(中国人民公安大学 信息技术与网络安全学院，北京 102628)

摘　要：如何提高特征点检测与匹配结果的精度，更好地优化相机位姿恢复结果，是提高三维重建整体效率的关键因素之一。基于 SIFT 算法原理，构建了一个全新的算法框架，该算法使用 FCN（fully convolutional networks，全卷积神经网络）神经网络和 BP（back propagation，反向传播）神经网络，综合考虑图像主目标的语义分割、图像灰度共生矩阵等方面的影响，实现了自适应的特征点检测范围、数量调整，并在特征点匹配阶段利用相机位姿偏移稳定性剔除误匹配，同时采用基于图优化的方法对位姿恢复结果进行非线性优化，得到了更加精确相机位姿。最后与现有的主流算法进行分析比对，实验结果验证了该算法的有效性，提高了特征点检测的场景自适应程度与特征点匹配、位姿恢复的精度，实现了更加高效的三维重建。

关键词：三维重建；特征点；人工神经网络；位姿恢复；非线性优化

中图分类号：TP391.41　　doi: 10.3969/j.issn.1001-3695.2018.05.0359

Feature point processing and position recovery optimization algorithm in 3D reconstruction system

Xu Jianpeng, Bu Fanliang†

(*School of Information Technology & Cyber Security*, *People's Public Security University of China*, *Beijing* 102628, *China*)

Abstract: How to improve the accuracy of feature point detection and matching results and to optimize the recovery results of camera pose is one of the key factors of the overall efficiency improvement of 3D reconstruction. In this paper, on the basis of the principle of SIFT algorithm, we constructed a completely new algorithm framework. The algorithm used FCN (Fully Convolutional Networks) neural network and BP (Back Propagation) neural network to comprehensively consider the semantic segmentation and image gray level co-occurrence matrix of the main target of image to achieve adaptive feature point detection range and quantity adjustment, and it used the offset stability of camera position to eliminate false matching during feature point matching. In the meantime, it optimized the pose recovery results by using graph-based optimization nonlinearity and obtained a more accurate camera pose. Finally, we compared it with the existing mainstream algorithms, and the experimental results verified the effectiveness of the proposed algorithm, the improvement of the scene-adaptive degree of feature points detection, the matching precision of feature points, the precision of posture recovery, and the better efficiency of three-dimensional reconstruction.

Key words: three-dimensional reconstruction; feature points; artificial neural networks; pose recovery; nonlinear optimization

0　引言

在三维重建的整个过程中，特征点的检测与匹配是最为基础的部分，其结果将作为对极几何、三角化等步骤的输入数据，以获得相机的位姿及空间点云的三维坐标，其中相机位姿是重建过程中的核心数据之一，其一方面反映了特征点处理阶段的结果，另一方面在后续的稠密点云恢复、纹理贴图等步骤中都将作为重要的数据输入源，直接对最终的三维重建效果产生决定性影响。因此，如何提高特征点检测、匹配及相机位姿的精度，是实现更加高效三维重建的关键所在。首先，在现有的主流特征点检测算法中，如 SIFT、SURF、ORB 等[1-3]，其检测范围覆盖整幅图像，而三维重建只关注于图像中的主要目标，如建筑、特定物体等，尤其是当主要目标之外的边缘点过多或存在相似像素块时，易产生误匹配[4]。其次，现有主流算法无法根据应用的需求来对特征点的数量进行控制，过多的特征点会导致大量的误匹配，影响位姿恢复的精度，而过少的特征点将无

收稿日期：2018-05-31；**修回日期：**2018-07-11　　**基金项目：**北京自然科学基金资助项目（4184099）；中国人民公安大学基础科研经费资助项目（2016JKF01316）

作者简介：徐建鹏（1993-），男，江西上饶人，硕士研究生，主要研究方向为信号与信息处理（cppsujianpeng@126.com）；卜凡亮（1965-），男（通信作者），教授，博士（后），主要研究方向为信号与信息处理（cppsujianpeng@126.com）.

法满足三维重建的数据量需求。最后，基于 K 最近邻（K-nearest neighbor，KNN）[5]、RANSAC（random sample consensus，随机抽样一致）[6]等主流算法的特征点匹配优化方法，其过程依赖于现有的匹配数据，即容易受误匹配的累积误差影响，优化结果不稳定，且其精度将直接影响到位姿的恢复。

人工神经网络在计算机视觉领域的快速发展为三维重建算法研究提供了一种新的途径， Hou 等[7]利用 caffe 框架下的 Alex Net 模型进行特征提取，特征描述鲁棒性能优于传统特征，且特征提取更加迅速；Mc Cormac 等[8]提出基于卷积神经网络的稠密 3 维语义地图构建方法 Semantic Fusion，利用卷积神经网络预测像素级的物体类别标签，最终生成包含语义信息的稠密 3 维语义地图；文献[9]通过卷积神经网络对图像进行自动分割，较好地解决三维重建中目标分割任务繁重的问题。另一方面，针对上述所分析的现存问题，文献[10,11]通过实际特征点数、图像对比度等因素来对相关的系数进行调整，实现特征点数量控制，取得了较好的效果，但由于考虑的因素较为单一，算法的自适应程度仍有待提高；文献[12]通过改进的 BRIEF 算子和对极几何原理来对特征点匹配进行优化，提高了匹配精度，但仍受制于误匹配的影响，算法结果的稳定性不足。主流的位姿优化方法首先需要恢复出三维点云，再利用重投影误差来进行非线性优化，但点云恢复过程中累积了位姿恢复阶段的误差，精度不足。

针对上述问题，本文基于 SIFT 算法原理提出了一个全新的算法框架。首先，本文通过构造符合三维重建需求的训练集数据，对 FCN 神经网络[13]进行训练，使其能够对图像中的主要目标进行检测，从而实现特征点检测范围的调整，同时对图像灰度共生矩阵进行分析，结合 SIFT 算法中的对比度阈值来构造 BP 神经网络的训练集数据，以拟合特征点数量、图像灰度共生矩阵[14]及对比度阈值之间的非线性关系，实现图像特征点数量的控制，其次，通过分析三角测量原理，利用其位姿偏移稳定性来对特征点匹配结果进行优化，即控制匹配特征点在图像中的像素位置范围，从而避免误匹配在优化过程中所产生的累积误差，实现精确匹配，最后利用图优化[15]方法，构造匹配特征点像素位置及相机位姿之间的最小二乘问题，来对位姿进行非线性优化，充分发挥了本文特征点匹配优化算法的精度优势。与 SIFT 算法进行分析比对后，实验结果验证了本文算法的有效性与高效性。

1　相关理论

1.1　神经网络

BP 神经网络是一种根据误差反向传播算法训练的多层前向神经网络，通过计算期望输出与实际输出之间的误差，由输出向输入逐层调整节点的参数与阈值，经过反复迭代，最终实现目标精度。FCN 神经网络主要用于图像语义分割，该网络相较于卷积神经网络，区别主要在于 FCN 的最后一层仍然为卷积层，用于上采样还原出像素的分类及位置信息。

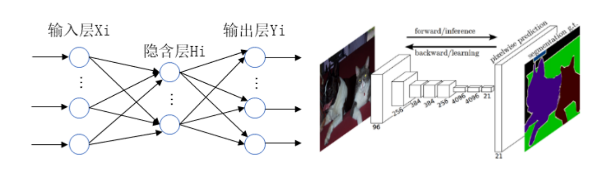


图 1　BP 神经网络与 FCN 神经网络结构

1.2　对极几何

如何恢复相机的空间运动是对极几何所解决的关键问题。如图 2 所示，对于三维空间中的一点 P，相机（光心位置）O_1，O_2 分别得到图像 I_1， I_2， P 在两幅图像中的像素分别表示为 p_1， p_2，光心连线 O_1O_2 称为基线，其与像素平面所产生的交点 e_1， e_2 称为极点， p_1， p_2 与 e_1， e_2 的连线 l_1， l_2 称为极线，且 O_2 的位置是由 O_1 经过空间变换 R ，t 得到。在特征点匹配阶段，本文已经获得了若干对相互匹配的特征点，即点 P 所产生的 p_1 和 p_2，若匹配关系未知，那么 p_1 所对应的匹配点可能在 l_2 的任何位置，如 $p_2^{'}$，由此所对应的空间坐标点则为 P'。因此，正确的特征点匹配是决定三维重建结果的最为重要的因素。

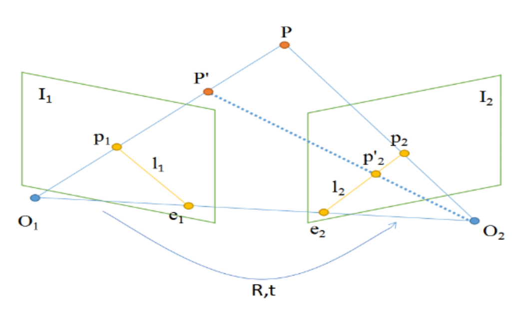


图 2　对极几何原理图

2　算法框架

根据以上分析，本文算法框架主要特征点检测、特征点匹配及位姿优化三个部分组成，具体如图 3 所示。

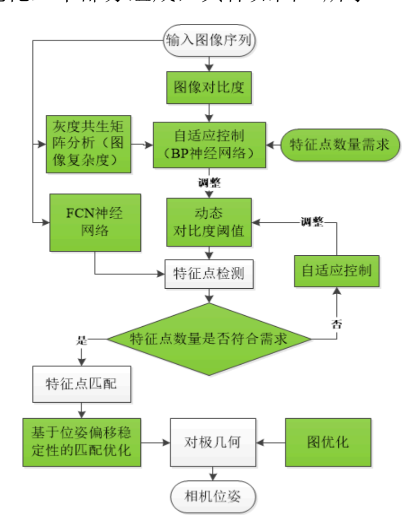


图 3　本文算法框架

在特征点检测阶段，该算法对输入的每幅图像计算其对比度与灰度共生矩阵，据此分析图像的复杂度，并明确应用所需的特征点数量，然后输入到基于 BP 神经网络的自适应控制模块中，由此调整对比度阈值，控制图像的整体特征点数量，若第一次调整后特征点数量仍不符合要求，将进行迭代调整，最

终使其符合需求。与此同时，图像还将经过训练好的 FCN 神经网络的处理，得到主目标像素范围，并以此来对特征点检测的区域进行限制，避免背景噪声信息所带来的累积误差。

在特征点匹配阶段，需要对三维重建中的相机位姿偏移稳定性进行分析。在相机移动的过程中，每次拍摄之间的位移不应过小，但同时要考虑到过大的位移会导致特征的消失或者场景外观的变化，具体如图 4 所示，其中 Δe 表示在相同的匹配误差下（$\Delta\theta$），不同的相机位移（t_1，t_2）所恢复的点云结果（P'）与真实结果（P）之间的误差，O_1，O_2 代表相机的光心位置。因此对于图像序列中相邻的两幅图像，其对应的匹配特征点的像素位置不应有过大的变化，由此性质，本文即可对匹配结果进行优化，剔除大量的误匹配，提高重建精度。

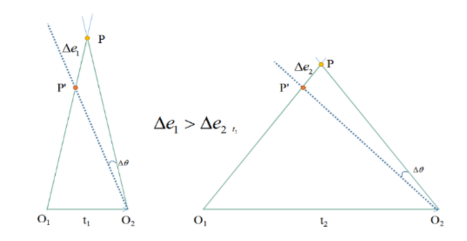


图 4　三维重建中相机位姿偏移稳定性原理

在相机位姿恢复阶段，至少需要五对匹配点即可恢复出相机的位姿 R 、 t，因此对于所有经过优化的匹配结果（通常都远大于 5），本文算法从概率论的角度进行分析，以图优化为求解工具，构造特征点像素坐标及相机位姿之间的最小二乘问题，即第一幅图像中的特征点，经过什么样位移变换，才最可能得到现有第二幅图像中匹配特征点的像素位置，由此来对相机位姿进行非线性优化，得到更符合实际的结果。

3　算法实现

3.1　BP 神经网络的构造与训练

对比度阈值是 SIFT 算法中用于提高特征点稳定性的关键参数之一，且不同的图像对比度会对特征点数量产生很大的影响。当本文需要定量地分析图像对比度时，本文一般采用均方根来进行计算，具体如式 1 所示，其中 n 表示像素数量，x_i 表示第 i 个像素的灰度值，$\bar{x}$ 表示图像像素灰度值的平均数[4]。

$$\mathrm{RMS}=\left[\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2\right]^{1/2} \tag{1}$$

对比度只是图像整体信息的一个方面，在相同对比度的图像下，复杂度也会极大地影响特征点数量。图像复杂度是对图像固有复杂性的描述和所有图像信息的聚集。为了获取定量的图像复杂度，本文通常使用灰度共生矩阵来分析图像的信息熵（information entropy, H）、相关度（correlation degree, COV）和能量（energy, J）等重要信息，具体如式 2 所示，其中 k 表示图像所具有的灰度级的数量，n_l 指第 l 个灰度级的总数量，N 代表图像的总像素数，$p(i,j)$ 表示灰度共生矩阵中第 (i,j) 个值，同时 $\mu_x=\sum_{x=1}^{k}\sum_{y=1}^{k}x\cdot p(x,y)$，$\mu_y=\sum_{x=1}^{k}\sum_{y=1}^{k}y\cdot p(x,y)$，$\sigma_x^2=\sum_{x=1}^{k}\sum_{y=1}^{k}p(x,y)\cdot(x-\mu_x)$，$\sigma_y^2=\sum_{x=1}^{k}\sum_{y=1}^{k}p(x,y)\cdot(y-\mu_y)$。

$$\begin{cases}\mathrm{H}=-\sum_{l=1}^{k}\frac{n_l}{N}log\left(\frac{n_l}{N}\right)\\ \mathrm{COV}=\left[\sum_{i=1}^{N}\sum_{j=1}^{N}ijp(i,j)-\mu_x\mu_y\right]/\sigma_x\sigma_y\\ \mathrm{J}=\sum_{i=1}^{N}\sum_{j=1}^{N}p(i,j)^2\end{cases} \tag{2}$$

通过以上分析，最终将信息熵、相关度、能量及边缘比例四个指标综合表示一幅图像的复杂度。求得图像复杂度的计算模型之后，本文需要构建特征点数、图像复杂度、图像 及对比度阈值之间的数学模型，以实现特征点数量的自适应控制。首先定义如式 3 所示的函数关系，该式可解释为，对于给定范围的特征点数量，模型自动匹配适当的对比度阈值，以使实际特征点数量与需求的相符。

$$\text{contrast threshold}=f(\text{complexity},\text{RMS},\text{point range}) \tag{3}$$

本文采用 BP 神经网络来拟合式 3 的非线性数学关系，从德国慕尼黑工业大学(Technische Universität München）数据集中共选取了 3000 张图像，对于每一张图像，构造了如表 1 所示的三组数据，由此共形成了包含 9000 组数据的训练集，其中的对比度阈值均通过反复的线性迭代来获得。

表 1　BP 神经网络训练集构造

图像序号	数据集	复杂度	对比度	特征点数量	对比度阈值
	1	0.4665	3.6986	900	0.08
1	2	0.4665	3.6986	800	0.12
	3	0.4665	3.6986	700	0.13

根据表 1，将复杂度（三个指标）、对比度及需求的特征点数量作为 BP 神经网络的输入，对比度阈值则作为输出。需要强调的是，选择 BP 神经网络的原因主要在于，相较于 LMS（least mean square）和 LM（Levenberg-Marquardt）等单层网络，BP 神经网络属于多层网络且具备更好的非线性拟合能力。本文使用 Matlab 工具构造出如图 5 所示的 BP 神经网络来拟合式 3 的数学模型，该神经网络的的隐含层共设置了 9 个节点，以提高泛化能力，输出设置 9 个节点，采用二进制（0,1）来表示对比度阈值，为确保神经网络的准确性，本文将训练误差精度设置为 0.251，同时对所有的输入进行了归一化处理，以消除量纲效应，经过如图 6 所示的 7424 次迭代后，神经网络达到目标精度，训练完成。最后，该神经网络将作为新 SIFT 框架中的自适应模块，根据图像的信息及应用需求，来自适应地调整对比度阈值，使最终检测结果符合预期。

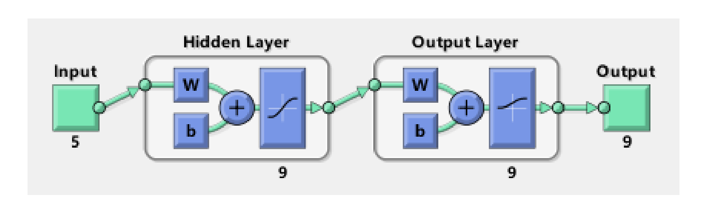


图 5　本文设计的 BP 神经网络结构

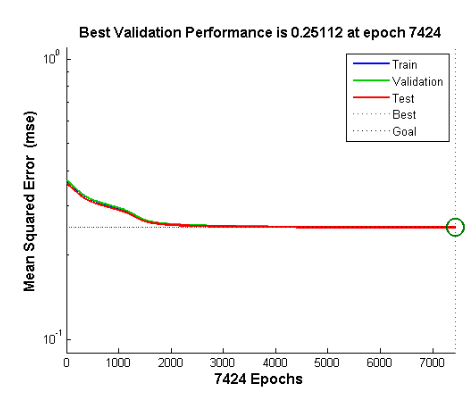


图 6　BP 神经网络训练过程

3.2　FCN 神经网络的构造与训练

要检测出图像中主目标的像素范围，传统的方法一般采用基于 Canny、Sobel 等算子的边缘检测算法对图像进行处理，但由于图像中还存在大量的背景图案，会产生大量的边缘信息，导致无法正确地计算出主目标的像素范围，因此本文采取 FCN 神经网络，进行像素级的分类，有效地提取出了场景中的主目标像素范围，提高了特征点检测的效率及匹配精度，具体网络结构如图 7 所示。

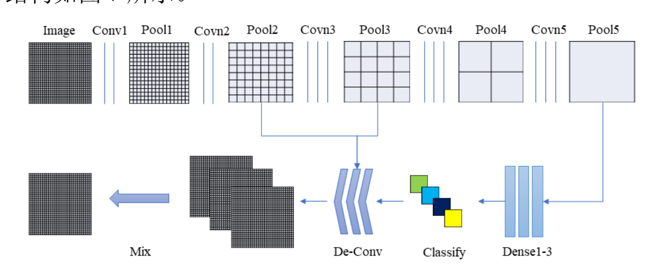


图 7　本文设计的 FCN 神经网络结构

该网络共包含五个卷积层（Conv1 与 Conv2 进行了两次卷积，其余进行了三次卷积）、五个池化层（pool）、三个全连接层（dense）和三个上采样层（de-conv），并将第二、第三池化层的输出结果作用到上采样过程中，以弥补在不断的下采样中所丢失的像素位置信息，提高分割的精度。本文通过 Caffe 深度学习框架来搭建图 7 中所设计的网络，并使用 PASCAL VOC2012 数据集作为该网络的训练集，最终将训练好的 FCN 网络应用于本文的算法框架中，用于提取图像中的主目标像素范围

3.3　基于位姿偏移稳定性的特征点匹配优化

根据上文对三维重建中相机位姿偏移稳定性的分析，设共有 n 对匹配点，在传统的 KNN 算法优化的基础上，对同时满足式 4 的匹配点进行保留， 即第二幅图像上对应的匹配点的位置，应与第一幅图像上匹配点位置保持一定范围内的约束，其中 Th_r 表示约束的范围，主要根据图像的像素大小来进行设置，一般取图像对角线像素长度的 10%。由此本文可以求得精确匹配点，并根据对极几何原理解出初始的 R、t 。

$$\begin{cases} M=\{((x_1,y_1),(x_1^{'},y_1^{'})),\cdots,((x_n,y_n),(x_n^{'},y_n^{'}))\} \\ (x_1^{'}-x_1)^2+(y_1^{'}-y_1)^2\le Th_r^2 \\ Th_r=\sqrt{(Image_col^2+Image_row^2)}\cdot 0.1 \end{cases} \tag{4}$$

3.4　基于图优化的相机位姿恢复

根据 1.3 中的原理，本文可以根据相机位姿优化问题设计如图 8 所示的图结构。

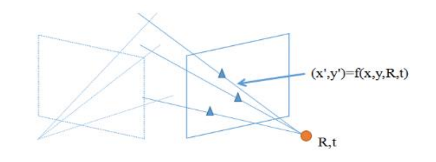


图 8　图优化结构

以上一个相机位姿为参考系，则当前相机位姿即可表示为(R，t)，即为图结构中的顶点（优化变量），图中的边（误差项）表示第一幅中的特征点，经过(R，t)旋转变换后，在第二幅图像中的像素位置，该像素位置也将作为误差项，与本文已得到的真实像素位置（即对应的一对匹配点中，第二幅图像中的点的像素位置）之间进行误差计算，由此本文即构建了一个以相机位姿为优化变量的最小二乘问题，最后可由基于图优化的第三方库 g2o(General Graphic Optimization)来具体实现与解决，具体分析如下。

首先构造式 5 所示的观测方程，表示在 R,t 相机位姿下观测其中 v_n 为符合高斯分布的噪声，$v_n\sim N(0,Q_n)$：

$$(x_n^{'},y_n^{'})=f((x_n,y_n),(R,t))+v_n \tag{5}$$

设状态变量 $w=\{(R,t),(x_1,y_1),\cdots,(x_n,y_n)\}$，$m=\{(x_1^{'},y_1^{'}),\cdots,(x_n^{'},y_n^{'})\}$，将原问题转换为构建最大化条件概率问题，如式 6 所示，即在已知第一幅图像及对应的第二幅图像的匹配点位置时，最可能的相机位姿是多少：

$$P(w|m)=\frac{P(m|w)P(w)}{P(m)}\propto P(m|w)P(w) \tag{6}$$

即求式 6 的最大化，其最大似然估计为

$$w_{MLE}^{*}=\arg\min\left((m_n-f((x_n,y_n),(R,t)))^T Q_n^{-1}(m_n-f((x_n,y_n),(R,t)))\right) \tag{7}$$

设误差项为

$$e_n=(x_n^{'},y_n^{'})-f((x_n,y_n),(R,t)) \tag{8}$$

由此可将式（6）转换为式（9）的形式：

$$\min E(n)=\sum_n e_n Q_N^{-1} e_n \tag{9}$$

在计算最小化 $E(n)$ 的过程中，若 $E(n)\ge Th_e$，则直接将其对应的匹配点对删除，一方面剔除了错误匹配，另一方面也减少了在求解 R、t 过程中所产生的累积误差，得到更加精确的结果。

4　实验与分析

实验硬件环境：Intel Core i7-6700HQ 四核处理器，8GB 内

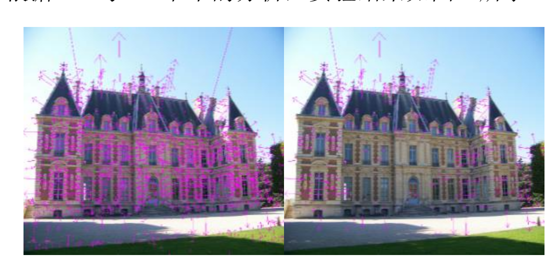
存，Linux Ubuntu14.04 操作系统，80GB 硬盘空间。

软件环境：VMware Workstation Pro，C++语言，g++编译器。

数据集：TUM 三维重建数据集，图像大小均为 255*255。

4.1 特征点检测

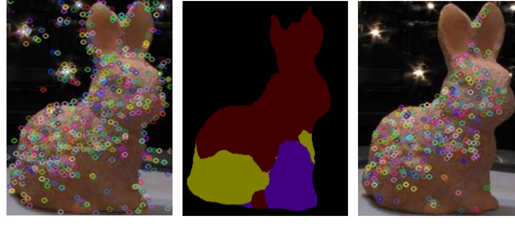
根据 3.1 与 3.2 节中的分析，实验结果如图 9 所示。



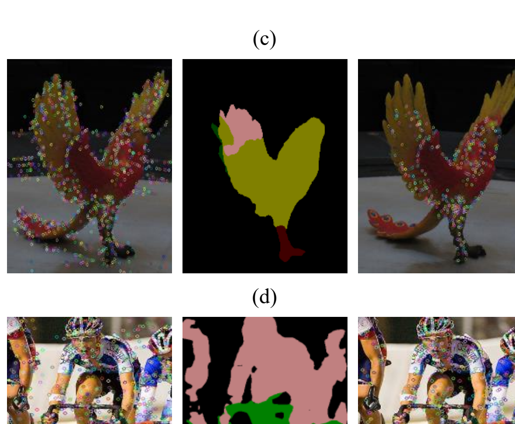
(a)



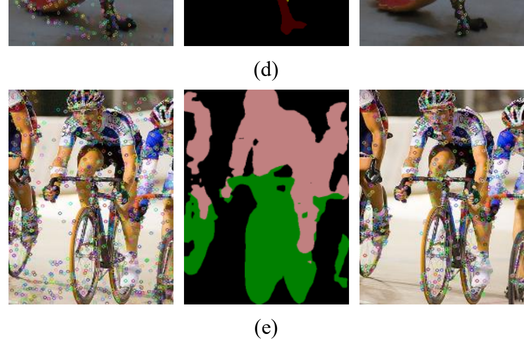
(b)



(c)



(d)



(e)

图 9　特征点检测实验结果图

在图 9(a)中，SIFT 算法共检测到 4 865 个特征点，经过本文算法调整后共检测到 1 084 个特征点，而对于图 9(b)中特征点过少的情况，本文算法将特征点数量由 231 调整到 973。图 9(c)~(e)左边的图像表示未经特征点检测范围控制的结果，在主目标外有大量的背景特征点，中间的图像表示 FCN 神经网络处理后的结果，提取出了主目标的范围，由此来对特征点进行范围控制，得到了集中于主目标上的特征点检测结果，如图 9(c)右图所示。相较于 SIFT 算法，本文算法在特征点检测阶段的耗时比较如表 2 所示。

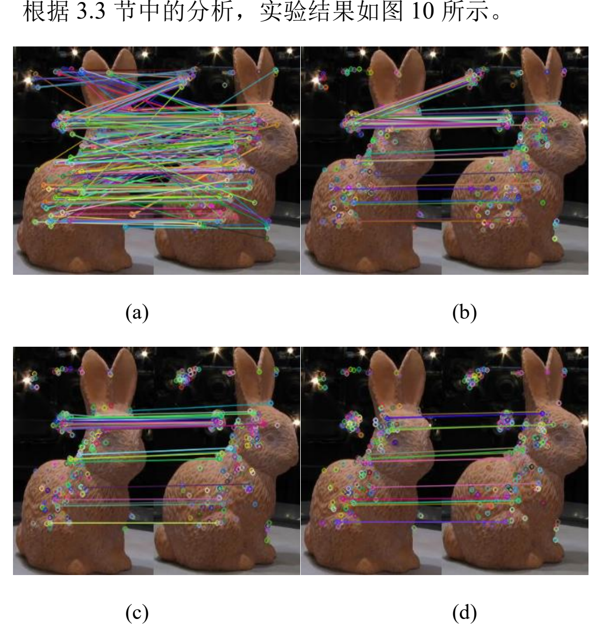
表 2　SIFT 与本文算法在特征点检测阶段的耗时比较结果

实验序号	1	2	3	4	5
SIFT 算法耗时	1.32	1.78	0.74	1.65	2.46
本文算法耗时	1.56	1.01	1.05	1.79	2.68

由以上实验结果可知，本文算法能够有效地对特征点数量进行调整，并实现了其检测范围的控制，其耗时较 SIFT 算法提高了 1.7%，其主要原因在于本文算法在 FCN 检测阶段耗费了更多的时间，但由于缩小了检测范围，因此总体上的耗时仍与 SIFT 算法持平。

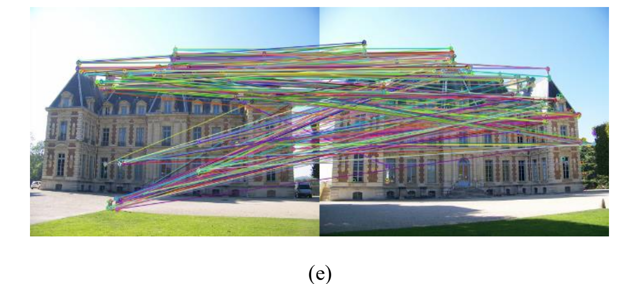
4.2 特征点匹配

根据 3.3 节中的分析，实验结果如图 10 所示。

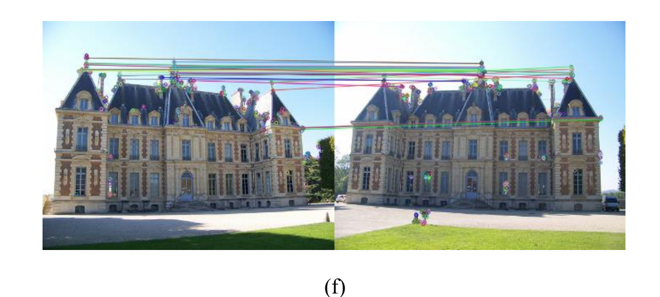


(a)　　(b)

(c)　　(d)



(e)



(f)

图 10　特征点匹配实验结果图

图 10(a)为未经任何优化的特征点匹配结果，图中存在大量的误匹配，图 10(b)表示经过传统的 KNN 算法优化的结果，匹配得到了一定的优化，但由于存在相似的像素块，导致仍存在明显的误匹配，经过基于位姿偏移稳定性的优化算法处理后，

得到图 10(c)所示的结果，再结合特征点检测范围的限制，即可得到图 10(d)，所有匹配均为精确匹配，类似的处理过程如图 10(e)(f)所示。相较于传统的基于 KNN 算法的匹配优化方式，其耗时比较如表 3 所示：

表 3　KNN 与本文算法在特征点匹配阶段的耗时比较结果

实验序号	1	2	3	4	5
KNN 算法耗时	2.33	2.69	1.14	2.45	3.02
本文算法耗时	2.42	2.75	1.23	2.61	3.24

由以上实验结果可知，基于位姿偏移稳定性的特征点优化算法能够有效地剔除误匹配，提高结果精度，为相机位姿恢复提供了准确的原始数据，其耗时较 KNN 算法提高了 5.3%，其主要原因在于增加了特征点像素位置计算等过程，但由于所增加的为线性计算，因此总体上的耗时仍与 KNN 算法基本持平。

4.3　相机位姿优化

为根据 3.4 节中的分析，设计如下所示伪代码来具体实现图优化方法，实验结果如图 11 所示。

输入：初始位姿 Rt_{init}，特征点匹配结果 M。

输出：优化后的位姿 Rt_{new}。

g2o( )

$\{Rt_{new} = Rt_{init}$

for(1)

for i in M :

$\text{error} = \text{error} + x_i * Rt - x_i^{'} + y_i * Rt - y_i^{'}$

if (error ≥ threshold) :

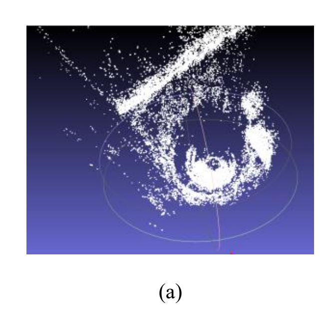
Delete(x_i, y_i)　else continue

if (error > $threshold^{'}$)

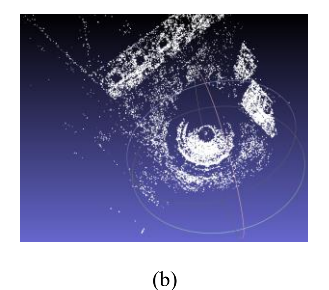
$$Rt_{new} = Rt_{new} - \propto \frac{\partial error}{\partial Rt_{new}}$$

error=0 and continue

else break and done



(a)



(b)

图 11　a）传统算法点云恢复结果；（b）本文算法点云恢复结果

根据未经优化的相机位姿所恢复的空间点云如图 11(a)所示，经本文算法优化后，其结果如图 11(b)所示，点云分布更加精确，主目标的细节轮廓更加明显。为定量分析其精度，本文使用重投影误差的方法来进行计算，即在现有的相机位姿视角下，观测空间点云，计算其投影到像素坐标系下的位置，再将此位置与实际对应的特征点像素坐标进行比较，分别统计其在 x 轴、y 轴的误差（以像素位单位计算），最终结果如表 4 所示。

表 4　传统算法与本文算法在点云恢复阶段的耗时及精度比较结果

传统算法	耗时	X 轴误差	Y 轴误差	本文算法	耗时	X 轴误差	Y 轴误差
1	0.10	1.45	4.26	1	0.20	0.15	0.0848
2	0.46	0.88	0.83	2	0.48	0.06	0.14
3	2.35	2.88	5.95	3	2.92	0.18	0.74
4	0.16	0.30	1.75	4	0.18	0.01	0.07
5	1.13	5.88	9.88	5	1.47	3.67	1.16

由以上实验结果可知，本文算法能够有效提高相机位姿的精度，相较于传统的优化算法，其 x 轴误差与 y 轴误差分别降低了 65%、91%，在耗时方面提高了 20%，其主要原因在于本文算法中，根据式 9，需要对误差大于阈值 Th_e 的匹配进行的剔除，增加了计算量。

4.4　综合分析

通过将上述三部分算法进行整合，本文对图 12 两幅图像进行了点云恢复处理，其结果如图 13 所示。



图 12　综合分析实验图像

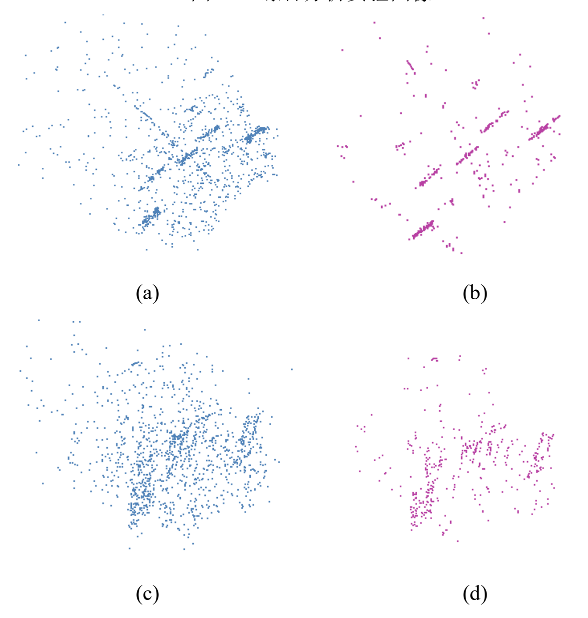


图 13　综合分析实验结果，(a)(b)为俯视，(c)(d)为侧视

图 13(a)(c)表示经传统主流算法（SIFT 特征点提取、KNN 特征点匹配优化、基于重投影误差的位姿优化）处理后的结果，共有 1302 个点云，但存在大量的错误数据，其重投影 x 轴误差为 16.21 个像素单位，y 轴误差为 35.56 个像素单位，算法耗时 5.52s。本文算法结果如图 14(b)(d)所示，共得到 446 个特征点，大量的错误点云数据得到了剔除，其重投影 x 轴误差为 0.16 个像素单位，y 轴误差为 5.18 个像素单位，算法耗时 5.86s。综上可知，本文算法通过自适应控制模块及 FCN 神经网络，在图像中的主目标区域提取了合适数量的特征点，剔除了大量的背景

特征点，且利用本文特征点匹配优化及位姿优化算法，得到了高精度的空间点云初始数据，由此提高三维重建结果的精度与效率。

5　结束语

本文针对目前特征点检测自适应程度不高，特征点匹配与位姿恢复精度不足等问题，提出了一个基于 SIFT 的全新算法框架，并对算法进行了实现。实验结果验证了本文算法的有效性，实现了更加高效的特征点检测、匹配及位姿恢复，为三维重建过程提供了高精度的空间点云数据，提升了整个三维重建过程的效率。但另一方面，本文算法的阈值 Th_r 应具备更好自适应程度，需根据后续所得的相机位姿来调整 Th_r 的取值，实现二次调整，以获得更多数量的精确匹配特征点，提高点云恢复的效率。

参考文献：

[1] Lowe D G. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision, 2004, 60 (2): 91-110.

[2] Bay H, Ess A, Tuytelaars T, *et al.* Speeded-up robust features [J]. Computer Vision & Image Understanding, 2008, 110 (3): 404-417.

[3] Rublee E, Rabaud V, Konolige K, *et al.* ORB: an efficient alternative to SIFT or SURF [C]// Proc of IEEE International Conference on Computer Vision. 2012: 2564-2571.

[4] 徐建鹏，卜凡亮. SIFT 动态对比度阈值算法 [J]. 中国人民公安大学学报：自然科学版, 2017, 23 (4): 61-64. (Xu Jianpeng, Bu Fanliang. Mining. SIFT dynamic contrast threshold algorithm [J]. Journal of People's Public Security University of China (Science and Technology) , 2017, 23 (4): 61-64.)

[5] Keller J M, Gray M R, Givens J A. A fuzzy K-nearest neighbor algorithm [J]. IEEE Trans on Systems Man & Cybernetics, 2012, SMC-15 (4): 580-585.

[6] Kim J, Lee M. Robust Lane Detection based on convolutional neural network and random sample consensus [M]// Neural Information Processing. Springer International Publishing, 2014: 454-461.

[7] Costante G, Mancini M, Valigi P, *et al.* Exploring representa-tion learning with CNNs for frame-to-frame ego-motion esti-mation [J]. IEEE Robotics and Automation Letters, 2016, 1 (1): 18-25.

[8] Mccormac J, Handa A, Davison A, *et al.* SemanticFusion: dense 3D semantic mapping with convolutional neural networks [C]// Proc of IEEE International Conference on Robotics and Automation. Piscataway, NJ: IEEE Press, 2017: 4628-4635.

[9] 温佩芝，苗渊渊，周迎，等. 基于卷积神经网络改进的图像自动分割方法 [J]. 计算机应用研究, 2018, 35 (9): 2848-2852. (Wen Peizhi, Miao Yuanyuan, Zhou Ying, *et al.* Improved image automatic segmentation method based on convolution neural network [J]. Application Research of Computers, 2018, 35 (9): 2848-2852.)

[10] 卫保国，张海曦. 一种二次阈值调整 SIFT 算法 [J]. 计算机科学, 2014, 41 (6): 282-286. (Wei Baoguo, Zhang Haixi. Twofold adjusted threshold SIFT [J]. Computers Science, 2014, 41 (6): 282-286.)

[11] 王强，李柏林，罗建桥，等. 一种局部二值模式图像特征点匹配算法 [J/OL]. 计算机应用研究, 2019, 36 (2) . [2018-01-19]. http://www. arocmag. com/article/02-2019-02-046. html. (Wang Qiang, Li Bailin, Luo Jianqiao, *et al.* Local binary pattern image feature point matching algorithm [J/OL]. Application Research of Computers, 2019, 36 (2) . [2018-01-19]. http://www. arocmag. com/article/02-2019-02-046. html.)

[12] 秦晓飞，皮军强，李峰. 基于极线约束的 ORB 特征匹配算法 [J]. 计算机应用研究, 2018, 35 (9): 2865-2868. (Qin Xiaofei, Pi Junqiang, Li Feng. ORB feature matching based on epipolar constraint [J]. Application Research of Computers, 2018, 35 (9): 2865-2868.)

[13] Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation [J]. IEEE Trans on Pattern Analysis & Machine Intelligence, 2014, 39 (4): 640-651.

[14] 高振宇，杨晓梅，龚剑明，等. 图像复杂度描述方法研究 [J]. 中国图象图形学报, 2010, 15 (1): 129-135. (Gao Zhenyu, Yang Xiaomei, Gong Jianming. Research on image complexity description methods [J]. Journal of Image and Graphics, 2010, 15 (1): 129-135.)

[15] Kümmerle R, Grisetti G, Strasdat H, *et al.* g2o: a general framework for graph optimization [C]// Proc of IEEE International Conference on Robotics and Automation. 2011: 3607-3613.